DOCUMENT RESUME

ED 174 641 TM 009 263

AUTHOR Yap, Kim Onn
TITLE Pretest-Posttest Correlation and Regression Models.
PUB DATE Apr 79
NOTE 33p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Correlation; Data Analysis; Educational Testing; Evaluation Methods; *Mathematical Models; *Multiple Regression Analysis; *Post Testing; *Pretesting; Pretests; *Program Effectiveness; Program Evaluation; Simulation; Statistical Bias
IDENTIFIERS Elementary Secondary Education Act Title I; *RMC Models

ABSTRACT

The accuracy with which regression models estimate treatment effects is dependent upon a number of conditions. The stability of the regression line (a function of sample size and correlation between pretest and posttest) is said to be the most important of these conditions. The utility of regression models is proportional to the size of the correlation between pretest and posttest. As the size of the correlation increases, the predicted posttest scores of the treatment group decreases. This produces a corresponding increase in the difference between predicted and observed scores. It is further stated that in compensatory education projects, factors which lower the correlation between pretest and posttest for low scoring students may invalidate the results. Given these condition, this study examined the impact of the correlation between pretest and posttest on the accuracy with which regression models estimate treatment effects in Title I evaluation. More specifically, the effect of the correlation between pretest and posttest on the estimation of treatment effects with regression models was studied, using simulated data. Conclusions regarding the importance of the pretest-posttest correlation varied, depending on whether more emphasis is placed upon unbiased estimates or efficiency. (Author/JKS)

PRETEST-POSTTEST CORRELATION AND REGRESSION MODELS

Kim Onn Yap

Northwest Regional Educational Laboratory

A paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, April 8-12, 1979

NWREL TAC

2/79
3372A

PRETEST-POSTTEST CORRELATION AND REGRESSION MODELS

Kim Onn Yap

Northwest Regional Educational Laboratory

INTRODUCTION

The accuracy with which regression models estimate treatment effects is dependent upon a number of conditions. Of these conditions, the stability of the regression line (a function of sample size and correlation between pretest and posttest) is said to be the most important (Tallmadge & Wood, 1976; Tallmadge & Horst, 1976; Horst, Tallmadge & Wood, 1975).

Many writers (e.g., Horst, Tallmadge & Wood, 1975) have stated that the utility of regression models is proportional to the size of the correlation between pretest and posttest. As the size of the correlation increases, the predicted posttest scores of the treatment group decreases. This produces a corresponding increase in the difference between predicted and observed scores. It is further stated (Tallmadge & Horst, 1976) that in compensatory education projects, factors which lower the correlation between pretest and posttest for low scoring students may invalidate the results.

Thus, according to these writers, in using regression models one should use a pretest which has a high correlation with the posttest. The higher the correlation, the lower the magnitude of regression to the mean.

On the other hand, it is also possible to argue that a high pretest-posttest correlation is not important. Since selection is based on the pretest, a high or low correlation should not affect our confidence in the no-treatment expectation. A high or low correlation merely reflects the goodness of the selection criteria. In the special case when the pretest-posttest correlation is zero, the regression models become equivalent to a control group design where students are randomly selected for treatment and control groups.

While high correlations between pretest and posttest are recommended by proponents of regression models, the size of such correlations of course could not be pre-determined. In some cases (as when teacher ratings or some composite measure is used as pretest and a standardized test is used as posttest) the correlation can be expected to be low.

The objective of the present study was to examine the impact of the correlation between pretest and posttest on the accuracy with which regression models estimate treatment effects in Title I evaluation. More specifically, the study was to provide an answer to the question: Does correlation between pretest and posttest make a difference in estimting treatment effects with regression models?

PROCEDURE

To study the impact of the size of pretest/posttest correlation on the accuracy with which regression models estimate treatment effects, data resembling those suited for analysis through these models were simulated. The rudiments of the simulation were as follows:

$$Y_{1}ij = Xij + Eij \quad (1)$$

$$Y_{2}ij = Xij + Gij + TEij + E'ij, \quad (2)$$

where $Y_{1}ij$ is the pretest score of student i in group j; $Y_{2}ij$ is the posttest score of student i in group j; Xij is the true achievement level of student i in group j at pretest; Gij is growth attributable to factors other than the treatment for student i in group j; TEij is the treatment effect for student i in group j; and Eij and E'ij are error terms.

For purposes of the simulation, it was assumed that the mean growth rates (Gij's) for the treatment and control groups are equal. In equation (2), TEij's were set to equal zero for students in the control group to indicate the absence of treatment effects.

The values of Xij, Gij, TEij, Eij, and E'ij were made up of random numbers provided by GAUSS (IBM, 1968), a computer subroutine which generates normally distributed random numbers. The relative size of Xij, Eij, and E'ij was adjusted by means of multipliers. For example, the values of a set of Xij, Eij and E'ij may be obtained as follows:

$$Xij = .7\ (N_{1})$$

$$Eij = .3\ (N_{2})$$

$$E'ij = .3\ (N_{4})$$

where the Ns are random numbers. Means and standard deviations for the random numbers were chosen in such a way that $Y_{1}ij$ and $Y_{2}ij$ would have approximately a mean of 50 and a standard deviation of 21.06, respectively, to correspond with the mean and standard deviation of Normal Curve Equivalents (NCEs). For example, in

$$Y_{1ij} = Xij + Eij, \text{ where}$$

$$Xij = .7\ (N_1), \text{ and}$$

$$Eij = .3\ (N_2),$$

both N_{1s} and N_{2s} were given a mean of 50 and a standard deviation of 27.65. This gave Y_1ij a mean of 50 and a standard deviation of

$$\sqrt{(.7)^2\ (27.65)^2 + (.3)^2\ (27.65)^2}$$

which simplifies to 21.06.

The same procedure was used to give Y_2ij a mean of 50 and a standard deviation of 21.06. Means and standard deviations for Gij and TEij were determined by providing the appropriate parameter values to subroutine GAUSS. Gij was set to have a mean of 10 and a standard deviation of 10 and TEij was set to have a mean of 7 and a standard deviation of 7. These means and standard deviations had been chosen to reflect what is most likely to occur in real-life situations in terms of NCE scores.

Negative values provided by GAUSS, which occurred on few occasions, were dropped, resulting in slightly higher means and lower standard deviations for the variables.

The pretest data (Y_1ij) were first simulated. The hypothetical cases in each data set were rank-ordered. A strict cut-off, located at the 25th percentile point, was used to assign cases to the treatment ($j = 1$) and control ($j = 2$) groups.

After the hypothetical cases had been assigned to Title I or control groups, posttest data (Y_2ij) were simulated by means of equation (2), adding growth (Gij) and treatment effects (TEij) to pretest scores of students in treatment groups and only growth (Gij) to pretest scores of students in control groups.

Three critical parameters were manipulated in the simulation (see Table 1). First, pretest reliability was varied from .84 to .69. Second, the size of correlation between pretest and posttest was varied from .75 to .50 and then to .25. Third, sample size was made to vary from 100 to 200.

Table 1 about here

The manipulation of data reliability was based on Gulliksen's (1950) idea that a reliability coefficient can be expressed as the ratio of true variance to total variance. This means that we could vary reliability by applying different multipliers to the random numbers which make up the values of variables. For example, in

$$Y_{1ij} = Xij + Eij, \text{ where}$$
$$Xij = .7\ (N_1),$$
$$Eij = .3\ (N_2),$$

and $N_{1}s$ and $N_{2}s$ are given the same variance, the reliability coefficient of Y_{1ij} is given by

$$r_{Y_1Y_1} = \frac{\text{Var } Xij}{\text{Var } Xij + \text{Var } Eij}.$$

Since multiplying a set of numbers by a constant increases the variance by the square of the constant and since $N_{1}s$ and $N_{2}s$ have the same variance, we have

$$r_{Y_1Y_2} = \frac{(.7)^2}{(.7)^2 + (.3)^2} = .84.$$

That is, the reliability of Y_{1ij} is .84. It could readily be verified that by changing the multipliers to .6 for N_{1s} and .4 for N_{2s} we will have lowered data reliability to .69. In the simulation, data sets with pretest reliability of .69 and .84 were created.

Correlations between Y_{1ij} and Y_{2ij} were controlled by means of the following formula:

$$R_{\infty\infty} = \frac{r_{1I}}{\sqrt{r_{11}\, r_{II}}}$$

Reported by Gulliksen (1950, p. 101), the formula gives the correlation between a test and a criterion when each is increased to infinite length to attain a reliability of unity. Given that

$$Y_{1ij} = Xij + Eij, \text{ and}$$

$$Y_{2ij} = Xij + Gij + E'ij,$$

the two variables share a common true score component in Xij, with $R_{\infty\infty}$ reaching unity when both Y_{1ij} and Y_{2ij} are made perfectly reliable. (For simulation purposes, variance given to Gij was treated as error variance.) It follows that $\sqrt{r_{11}\, r_{II}} = r_{1I}$, which provides a means of obtaining a desired value for r_{1I} by changing r_{11}, r_{II}, or both.

In the present simulation we had required that r_{11} (reliability of Y_{1ij}) be either .84 or .69 (a fixed value), leaving r_{II} (reliability of Y_{2ij}) to be varied to yield a desired value for r_{1I}. The way in which a desired correlation, say .75, between Y_{1ij} and Y_{2ij} was obtained is illustrated as follows:

Since (a) $\sqrt{r_{11}\, r_{II}} = r_{1I}$,

(b) the desired value of r_{1I} was .75 and (c) r_{11} had been given a reliability of .84, we had

$$\sqrt{.84\, r_{II}} = .75$$
$$\sqrt{r_{II}} = \frac{.75}{\sqrt{.84}}$$
$$= .82$$
$$r_{II} = .67$$

In other words, giving Y_{2ij} a reliability of .67 produced a correlation of .75 between Y_{1ij} and Y_{2ij}.

Since the variance of Y_{2ij} was made to equal 443.52 (the square of 21.06), the true variance required to yield a reliability coefficient of .67 was (443.52)×(.67) which equals 297.16. An appropriate multiplier (.62 in this case) was then applied to Xij in Y_{2ij} (Xij had a standard deviation of 27.65 when pretest reliability was .84) to produce the required true variance.

To obtain correlation coefficients of .75, .50 and .25 between Y_{1ij} and Y_{2ij}, the reliability of Y_{2ij} was varied from .67 to .07 for data sets where pretest (Y_{1ij}) reliability was .84 and from .78 to .09 for data sets where pretest reliability was .69. However, as indicated earlier, variance due to Gij was treated as error variance. Thus, the "reliability" coefficient for Y_{2ij} was more a measure of the amount of covariance between Y_{1ij} and Y_{2ij} than the ratio between true and total variance in Y_{2ij} itself. When variance due to Gij was treated as part of the true variance, the reliability of Y_{2ij} was found to range from .30 to .90 for data sets where pretest reliability was .84 and from .32 to .99 for data sets where pretest reliability was .69.

The correlations between Y_{1ij} and Y_{2ij} were determined before treatment effect (TEij) was added as a component to Y_{2ij} for Title I groups. Since TEij was given a population standard deviation of 7 (or a variance of 49), the correlations between Y_{1ij} and Y_{2ij} for Title I groups would differ slightly from the correlations for the comparison groups.

THE DATA SETS

Taking into account the different levels of each of the three parameters (i.e., size of correlation between pretest and posttest, data reliability, sample size), a total of 3 x 2 x 2 (or 12) categories of data sets were simulated. One hundred data sets were created for each of the categories. Characteristics of these data sets are summarized in Appendices A to L.

An examination of the characteristics of these data sets suggests that they closely resemble what we had intended to create. The obtained values, in some instances, deviate slightly from the parameters. As explained earlier, this came about essentially as a result of dropping negative values provided by GAUSS on a few occasions. Except for the slightly higher means and lower standard deviations, the data have the appearance of NCE scores. (The higher means for Y_{2ij} are due to higher means for Xij and Gij.)

ANALYSIS AND RESULTS

Data analysis procedures described by Tallmadge and Wood (1976, 1978) for regression models were applied to the simulated data. Essentially, the procedures require that a regression line be determined on the basis of comparison group data to predict what the performance of Title I group would have been--if they had not received Title I treatment. The prediction is made at the point where the Title I group's pretest mean intercepts the regression line. The predicted performance is then subtracted from the actual performance of the Title I group, with the remainder being the estimated treatment effect or gain.

In the simulation this estimated gain was again subtracted from the actual gain (TEij) which was built into the posttest (Y_{2ij}) of the Title I group. The difference was then interpreted as an index of the accuracy with which regression models estimate treatment effects. The means and standard deviations of such differences for each of the 12 data categories are summarized in Table 2.

Table 2 about here

DISCUSSION

Before we examine the effects which the manipulated parameters (i.e., pretest-posttest correlation, pretest reliability, sample size) have on the estimation of treatment effects it might be helpful to present a perspective in which the results will be interpreted. As Wonnacott and

Wonnacott (1970) point out, an estimator may be described in terms of bias, efficiency and consistency. An unbiased estimator is one that is, on the average, right on target. In other words, its expected value is identical with the true value of the parameter. A biased estimator, on the other hand, has an expected value that is "off target" or deviates from the true value of the parameter. An efficient estimator is an unbiased estimator with a relatively small variance. An inefficient estimator, on the other hand, is an unbiased estimator with a relatively large variance. A consistent estimator is one which zeroes in on the true value of the parameter as sample size increases.

Viewed in this perspective, the results in Table 2 suggest that regression models can be depended upon to provide relatively unbiased estimates of treatment effects. This is so because the mean differences between the estimated and actual gains were in general negligibly small. Only in one case (Category V) did the mean difference exceed an absolute value of 1.0. It is interesting to note that none of the parameters being manipulated in the simulation appeared to have any appreciable or systematic effects on the amount of bias in estimation.

However, with respect to efficiency we have a different story. The results show that the size of correlation between pretest and posttest has a clearly discernible effect on the efficiency of the estimates. This is reflected in the increasingly larger standard deviations for the mean differences as we move from a higher correlation to a lower correlation. This pattern is seen across all 12 categories of data sets. For the first three categories (where the pretest has a reliability of .84) the standard deviation of the mean differences increases from 3.53 to 4.93 as the pretest-posttest correlation decreases from .78 to .25. For the next three categories (where the pretest reliability is lowered to .69) the increase in standard deviation is even greater, rising from 4.47 to 6.18 as the pretest-posttest correlation drops from .75 to .27. Similar patterns are seen in data sets in the other categories.

Sample size also seems to have some effect on the distribution of the mean differences or the efficiency of the estimates. This is evidenced by the fact that standard deviations for data categories VII to XII (where sample size is 200) generally are smaller than those for the first six categories (where sample size is 100). The former standard deviations range from 2.88 to 4.59, the latter from 3.53 to 6.18. It would appear that a smaller sample size serves to further reduce the efficiency of the estimates.

That regression models do not provide us with efficient estimates of treatment effects is perhaps more dramatically shown by the ranges of the mean differences obtained for the various data categories. These are reported in Table 3. As one can hardly fail to be struck by the magnitude of the ranges, one must remember that these are the extremes. While they could occur and did occur in the simulation, the probability of their occurring in any single evaluation is low. The standard deviations of mean differences discussed earlier remain a more stable measure of the efficiency of the estimates.

Table 3 about here

As indicated earlier, an increase in sample size appears to enhance the efficiency of the estimates. A larger sample size, however, does not appear to produce a more consistent estimator--at least not in any systematic way. It is true that in absolute value the sum of mean differences obtained for data categories VII to XII (where sample size is 200) is less than that obtained for data categories I to VI (where sample size is 100), the former being 1.75 and the latter 2.2. It is also true, however, that in absolute value the mean differences for data categories I and II are smaller than the mean differences for data categories VII and VIII. This is also true with respect to data categories IV and VI and data categories X and XII. It would appear that with regression models a larger sample size does not necessarily result in an increase in the consistency of estimates.

CONCLUDING REMARKS

In the simulation the author had set out to obtain evidence which would support one of the two arguments concerning the importance of having a high pretest-posttest correlation when regression models are used to estimate treatment effects. Quite ironically, the evidence, as it turned out, appeared to support both arguments in different ways. If one used bias as the sole criterion to judge the adequacy of an estimator, the evidence would suggest that the size of pretest-posttest correlation appears to have neither appreciable nor systematic effects on the estimation of treatment effects. In general, regression models provide estimates that are practically unbiased.

This does not mean, however, that pretest-posttest correlation makes no difference whatsoever. For if one used efficiency as a criterion to assess the adequacy of an estimator, the evidence would suggest that pretest-posttest correlation has appreciable and systematic effects on the estimation of treatment effects. The pattern of results clearly indicates that the lower the correlation the less efficient the estimates. In addition, sample size also appears to have some effect on efficiency, with larger sample sizes tending to enhance the efficiency of the estimates.

Therefore, in the case of a single evaluation—which is typically the case in a Title I district—the effects of pretest-posttest correlation on evaluation results can by no means be ignored. It would not be difficult to see why upon a perusal of the ranges of mean differences reported in Table 3 a Title I project administrator would seek a higher pretest-posttest correlation or a different evaluation model to estimate treatment effects.

Perhaps the most significant findings of the present study relate to the overall adequacy of regression models when used to estimate Title I treatment effects. In a typical situation, the pretest-posttest correlation derived from Title I evaluation data would fall within the range of those simulated in the present study (i.e., from .25 to .75). The results of the simulation suggest that within that range of correlation coefficients, regression models provide relatively inefficient—although relatively unbiased—estimates of treatment effects. The confidence intervals (as indicated by the standard deviations of mean differences) are clearly too large to be sensitive to small achievement gains typically produced in Title I projects. It would appear that the correlation coefficient of .40 recommended by Tallmadge and Wood (1978, p. 77) as, by rule of thumb, being sufficiently high to ensure reasonable accuracy of estimates of treatment effects, may, after all, be too liberal. The results of the present study would suggest a correlation coefficient of .75 or higher.

REFERENCES

Gulliksen, H. Theory of mental tests. New York: Wiley and Sons, 1950.

Horst, D. P., Tallmadge, G. K. & Wood, C. T. A practical guide to measuring project impact on student achievement. Mountain View, California: RMC Research Corporation, 1975.

IBM. System/360 scientific subroutine package. New York: IBM, 1968.

Tallmadge, G. K. & Horst, D. P. A procedural guide for validating achievement gains in educational projects. Mountain View, California: RMC Research Corporation, 1976.

Tallmadge, G. K. & Wood, C. T. Users guide: ESEA Title I evaluation and reporting system. Mountain View, California: RMC Research Corporation, 1976.

Tallmadge, G. K. & Wood, C. T. Users guide: ESEA Title I evaluation and reporting system (Revised). Mountain View, California: RMC Research Corporation, 1978.

Wonnacott, R.J. & Wonnacott, T. H. Econometrics. New York: Wiley and Sons, 1970.

Table 1

Parameters manipulated in the Simulation Study

Parameter	Level
Size of Pretest-Posttest Correlation	.75, .50, .25
Pretest Reliability	.84, .69
Sample Size	100, 200

Table 2

Differences Between Actual and Estimated Gains by Data Set Category and Size of Pretest-Posttest Correlation

Data Set Category	Pretest-Posttest Correlation	Estimated Gain	Actual Gain	Difference	
				Mean	S.D.
I	.78	8.82	8.89	.08	3.53
II	.54	9.15	8.86	-.29	4.88
III	.25	8.45	8.93	.48	4.93
IV	.75	8.81	8.89	.09	4.47
V	.53	9.91	8.86	-1.05	5.28
VI	.27	9.14	8.93	-.21	6.18
VII	.78	9.07	8.88	-.19	2.88
VIII	.54	8.69	9.05	.36	3.83
IX	.24	8.81	8.97	.16	4.78
X	.75	9.06	8.88	-.18	2.69
XI	.52	9.01	9.05	.04	4.07
XII	.26	8.14	8.97	.82	4.59

Table 3

Range of Differences Between Actual and Estimated Gains by Data Set Category and Size of Pretest-Posttest Correlation

Data Set Category	Pretest-Posttest Correlation	Range of Differences	
		From	To
I	.78	-7.51	8.02
II	.54	-12.24	13.37
III	.25	-8.63	14.86
IV	.75	-9.38	9.33
V	.53	-16.89	13.45
VI	.27	-13.48	12.37
VII	.78	-6.95	6.53
VIII	.54	-7.64	9.59
IX	.24	-12.66	13.04
X	.75	-6.63	6.46
XI	.52	-10.87	8.69
XII	.26	-12.46	13.62

Footnotes for Appendices A-L

1. The notations in the Appendices are interpreted as follows:

   r_{xx} = pretest reliability

   $ry_{1}y_{2}$ = correlation between pretest and posttest

   $\bar{Y}_1$ = pretest mean

   Sy_1 = pretest standard deviation

   $\bar{Y}_2$ = posttest mean

   Sy_2 = posttest standard deviation

   $\bar{G}$ = growth mean

   Sg = growth standard deviation

2. Each data category consists of 100 data sets. For categories I-VI, each of the 100 data sets consists of 100 simulated cases. For categories VII-XII, each of the 100 data sets consists of 200 simulated cases.

3. S.D. in the last column refers to standard deviations for the 100 simulated data sets.

Appendix A

Characteristics of Data Sets in Category I

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{y_1y_2}$	.78	.04
$\overline{Y}_1$	52.33	1.73
S_{y_1}	19.38	1.19
$\overline{Y}_2$	55.76	1.75
S_{y_2}	18.43	1.17
$\overline{G}$	12.91	.83
S_g	7.91	.55

Appendix B

Characteristics of Data Sets in Category II

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{Y_1Y_2}$	.54	.08
$\bar{Y}_1$	52.10	1.95
S_{Y_1}	19.42	1.30
$\bar{Y}_2$	56.08	1.76
S_{Y_2}	17.89	1.17
$\bar{G}$	12.86	.75
S_g	7.96	.57

Appendix C

Characteristics of Data Sets in Category III

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{Y_1Y_2}$	.25	.09
$\bar{Y}_1$	52.42	1.75
S_{Y_1}	19.53	1.27
$\bar{Y}_2$	54.69	1.76
S_{Y_2}	18.28	1.28
$\bar{G}$	12.78	.89
S_g	7.83	.53

Appendix D

Characteristics of Data Sets in Category IV

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{Y_1Y_2}$	.75	.04
$\bar{Y}_1$	52.75	1.87
S_{Y_1}	19.13	1.35
$\bar{Y}_2$	54.58	2.04
S_{Y_2}	18.86	1.47
$\bar{G}$	12.82	.83
S_g	7.86	.52

Appendix E

Characteristics of Data Sets in Category V

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{y_1y_2}$	.53	.07
$\bar{Y}_1$	52.64	1.89
sy_1	19.15	1.35
$\bar{Y}_2$	56.35	1.78
sy_2	17.91	1.04
$\bar{G}$	12.87	.85
sg	7.92	.56

Appendix F

Characteristics of Data Sets in Category VI

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{Y_1Y_2}$	.27	.10
$\overline{Y}_1$	52.49	1.80
S_{Y_1}	19.21	1.19
$\overline{Y}_2$	55.30	2.08
S_{Y_2}	18.77	1.16
$\overline{G}$	12.79	.68
S_g	7.84	.57

Appendix G

Characteristics of Data Sets in Category VII

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{y_1 y_2}$	.78	.03
$\bar{Y}_1$	52.16	1.35
S_{y_1}	19.35	.91
$\bar{Y}_2$	55.52	1.27
S_{y_2}	18.59	.93
$\bar{G}$	12.92	.60
S_g	7.91	.37

Appendix H

Characteristics of Data Sets in Category VIII

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{Y_1Y_2}$	.54	.05
$\bar{Y}_1$	51.86	1.30
S_{Y_1}	19.29	.98
$\bar{Y}_2$	56.42	1.40
S_{Y_2}	18.19	.84
$\bar{G}$	12.96	.56
S_g	7.94	.39

Appendix I

Characteristics of Data Sets in Category IX

Characteristics	Mean	S.D.
r_{xx}	.84	—
$r_{y_1 y_2}$	.24	.07
$\overline{Y}_1$	52.04	1.40
s_{y_1}	19.23	1.01
$\overline{Y}_2$	54.96	1.14
s_{y_2}	18.57	.92
$\overline{G}$	13.00	.51
s_g	7.95	.35

Appendix J

Characteristics of Data Sets in Category X

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{Y_1Y_2}$	.75	.03
$\overline{Y}_1$	52.61	1.25
S_{Y_1}	19.20	.83
$\overline{Y}_2$	54.52	1.42
S_{Y_2}	18.89	.78
$\overline{G}$	12.92	.55
S_g	7.94	.39

Appendix K

Characteristics of Data Sets in Category XI

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{Y_1Y_2}$	.52	.05
$\overline{Y}_1$	52.54	1.27
S_{Y_1}	19.09	.87
$\overline{Y}_2$	56.21	1.24
S_{Y_2}	18.04	.87
$\overline{G}$	13.03	.47
S_g	7.98	.34

Appendix L

Characteristics of Data Sets in Category XII

Characteristics	Mean	S.D.
r_{xx}	.69	—
$r_{y_1 y_2}$	.26	.07
$\bar{Y}_1$	52.67	1.41
Sy_1	19.07	.96
$\bar{Y}_2$	55.25	1.12
Sy_2	18.67	.91
$\bar{G}$	12.84	.59
Sg	7.93	.41